AU 1814

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/147,912*   #23   DATE: 01/07/94
TIME: 14:49:26

INPUT SET: S811.raw

18r Cu

```
 1                        SEQUENCE LISTING
 2
 3   (1)    General Information:
 4
 5      (i) APPLICANT: WAHL, DR., GEOFFREY M.
 6                    O'GORMAN DR., STEPHEN V.
 7
 8     (ii) TITLE OF INVENTION: FLP-MEDIATED GENE MODIFICATION IN
 9              MAMMALIAN CELLS, AND COMPOSITIONS AND CELLS USEFUL
10              THEREFOR
11
12    (iii) NUMBER OF SEQUENCES: 4
13
14     (iv) CORRESPONDENCE ADDRESS:
15              (A) ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
16              (B) STREET: 444 South Flower Street, Suite 2000
17              (C) CITY: Los Angeles
18              (D) STATE: CA
19              (E) COUNTRY: USA
20              (F) ZIP: 90071
21
22      (v) COMPUTER READABLE FORM:
23              (A) MEDIUM TYPE: Floppy disk
24              (B) COMPUTER: IBM PC compatible
25              (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26              (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
27
28     (vi) CURRENT APPLICATION DATA:
29              (A) APPLICATION NUMBER: US/08/147,912
30              (B) FILING DATE:
31              (C) CLASSIFICATION:
32
33    (vii) PRIOR APPLICATION DATA:
34              (A) APPLICATION NUMBER: US 07/666,252
35              (B) FILING DATE: 08-MAR-1991
36
37   (viii) ATTORNEY/AGENT INFORMATION:
38              (A) NAME: REITER MR., STEPHEN E.
39              (B) REGISTRATION NUMBER: 31192
40              (C) REFERENCE/DOCKET NUMBER: P31 8929
41
42     (ix) TELECOMMUNICATION INFORMATION:
43              (A) TELEPHONE: (619) 535-9001
44              (B) TELEFAX: (619) 535-8949
45
46
47   (2) INFORMATION FOR SEQ ID NO:1:
48
49      (i) SEQUENCE CHARACTERISTICS:
50              (A) LENGTH: 1380 base pairs
51              (B) TYPE: nucleic acid
```

ENTERED

INPUT SET: S811.raw

```
 52                   (C) STRANDEDNESS: single
 53                   (D) TOPOLOGY: linear
 54
 55        (ii) MOLECULE TYPE: DNA (genomic)
 56
 57
 58       (vii) IMMEDIATE SOURCE:
 59                   (B) CLONE: NATIVE FLP
 60
 61        (ix) FEATURE:
 62                   (A) NAME/KEY: CDS
 63                   (B) LOCATION: 1..1269
 64
 65
 66        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
 67
 68   ATG CCA CAA TTT GAT ATA TTA TGT AAA ACA CCA CCT AAG GTG CTT GTT        48
 69   Met Pro Gln Phe Asp Ile Leu Cys Lys Thr Pro Pro Lys Val Leu Val
 70    1               5                  10                  15
 71
 72   CGT CAG TTT GTG GAA AGG TTT GAA AGA CCT TCA GGT GAG AAA ATA GCA        96
 73   Arg Gln Phe Val Glu Arg Phe Glu Arg Pro Ser Gly Glu Lys Ile Ala
 74                  20                  25                  30
 75
 76   TTA TGT GCT GCT GAA CTA ACC TAT TTA TGT TGG ATG ATT ACA CAT AAC       144
 77   Leu Cys Ala Ala Glu Leu Thr Tyr Leu Cys Trp Met Ile Thr His Asn
 78              35                  40                  45
 79
 80   GGA ACA GCA ATC AAG AGA GCC ACA TTC ATG AGC TAT AAT ACT ATC ATA       192
 81   Gly Thr Ala Ile Lys Arg Ala Thr Phe Met Ser Tyr Asn Thr Ile Ile
 82          50                  55                  60
 83
 84   AGC AAT TCG CTG AGT TTC GAT ATT GTC AAT AAA TCA CTC CAG TTT AAA       240
 85   Ser Asn Ser Leu Ser Phe Asp Ile Val Asn Lys Ser Leu Gln Phe Lys
 86    65                  70                  75                  80
 87
 88   TAC AAG ACG CAA AAA GCA ACA ATT CTG GAA GCC TCA TTA AAG AAA TTG       288
 89   Tyr Lys Thr Gln Lys Ala Thr Ile Leu Glu Ala Ser Leu Lys Lys Leu
 90                  85                  90                  95
 91
 92   ATT CCT GCT TGG GAA TTT ACA ATT ATT CCT TAC TAT GGA CAA AAA CAT       336
 93   Ile Pro Ala Trp Glu Phe Thr Ile Ile Pro Tyr Tyr Gly Gln Lys His
 94                 100                 105                 110
 95
 96   CAA TCT GAT ATC ACT GAT ATT GTA AGT AGT TTG CAA TTA CAG TTC GAA       384
 97   Gln Ser Asp Ile Thr Asp Ile Val Ser Ser Leu Gln Leu Gln Phe Glu
 98              115                 120                 125
 99
100   TCA TCG GAA GAA GCA GAT AAG GGA AAT AGC CAC AGT AAA AAA ATG CTT       432
101   Ser Ser Glu Glu Ala Asp Lys Gly Asn Ser His Ser Lys Lys Met Leu
102              130                 135                 140
```

INPUT SET: S811.raw

```
103
104    AAA GCA CTT CTA AGT GAG GGT GAA AGC ATC TGG GAG ATC ACT GAG AAA        480
105    Lys Ala Leu Leu Ser Glu Gly Glu Ser Ile Trp Glu Ile Thr Glu Lys
106    145                 150                 155                 160
107
108    ATA CTA AAT TCG TTT GAG TAT ACT TCG AGA TTT ACA AAA ACA AAA ACT        528
109    Ile Leu Asn Ser Phe Glu Tyr Thr Ser Arg Phe Thr Lys Thr Lys Thr
110                        165                 170                 175
111
112    TTA TAC CAA TTC CTC TTC CTA GCT ACT TTC ATC AAT TGT GGA AGA TTC        576
113    Leu Tyr Gln Phe Leu Phe Leu Ala Thr Phe Ile Asn Cys Gly Arg Phe
114                        180                 185                 190
115
116    AGC GAT ATT AAG AAC GTT GAT CCG AAA TCA TTT AAA TTA GTC CAA AAT        624
117    Ser Asp Ile Lys Asn Val Asp Pro Lys Ser Phe Lys Leu Val Gln Asn
118                    195                 200                 205
119
120    AAG TAT CTG GGA GTA ATA ATC CAG TGT TTA GTG ACA GAG ACA AAG ACA        672
121    Lys Tyr Leu Gly Val Ile Ile Gln Cys Leu Val Thr Glu Thr Lys Thr
122        210                 215                 220
123
124    AGC GTT AGT AGG CAC ATA TAC TTC TTT AGC GCA AGG GGT AGG ATC GAT        720
125    Ser Val Ser Arg His Ile Tyr Phe Phe Ser Ala Arg Gly Arg Ile Asp
126    225                 230                 235                 240
127
128    CCA CTT GTA TAT TTG GAT GAA TTT TTG AGG AAT TCT GAA CCA GTC CTA        768
129    Pro Leu Val Tyr Leu Asp Glu Phe Leu Arg Asn Ser Glu Pro Val Leu
130                        245                 250                 255
131
132    AAA CGA GTA AAT AGG ACC GGC AAT TCT TCA AGC AAT AAA CAG GAA TAC        816
133    Lys Arg Val Asn Arg Thr Gly Asn Ser Ser Ser Asn Lys Gln Glu Tyr
134                    260                 265                 270
135
136    CAA TTA TTA AAA GAT AAC TTA GTC AGA TCG TAC AAT AAA GCT TTG AAG        864
137    Gln Leu Leu Lys Asp Asn Leu Val Arg Ser Tyr Asn Lys Ala Leu Lys
138        275                 280                 285
139
140    AAA AAT GCG CCT TAT TCA ATC TTT GCT ATA AAA AAT GGC CCA AAA TCT        912
141    Lys Asn Ala Pro Tyr Ser Ile Phe Ala Ile Lys Asn Gly Pro Lys Ser
142        290                 295                 300
143
144    CAC ATT GGA AGA CAT TTG ATG ACC TCA TTT CTT TCA ATG AAG GGC CTA        960
145    His Ile Gly Arg His Leu Met Thr Ser Phe Leu Ser Met Lys Gly Leu
146    305                 310                 315                 320
147
148    ACG GAG TTG ACT AAT GTT GTG GGA AAT TGG AGC GAT AAG CGT GCT TCT       1008
149    Thr Glu Leu Thr Asn Val Val Gly Asn Trp Ser Asp Lys Arg Ala Ser
150                    325                 330                 335
151
152    GCC GTG GCC AGG ACA ACG TAT ACT CAT CAG ATA ACA GCA ATA CCT GAT       1056
153    Ala Val Ala Arg Thr Thr Tyr Thr His Gln Ile Thr Ala Ile Pro Asp
```

INPUT SET: S811.raw

```
154              340              345              350
155
156  CAC TAC TTC GCA CTA GTT TCT CGG TAC TAT GCA TAT GAT CCA ATA TCA    1104
157  His Tyr Phe Ala Leu Val Ser Arg Tyr Tyr Ala Tyr Asp Pro Ile Ser
158          355              360              365
159
160  AAG GAA ATG ATA GCA TTG AAG GAT GAG ACT AAT CCA ATT GAG GAG TGG    1152
161  Lys Glu Met Ile Ala Leu Lys Asp Glu Thr Asn Pro Ile Glu Glu Trp
162      370              375              380
163
164  CAG CAT ATA GAA CAG CTA AAG GGT AGT GCT GAA GGA AGC ATA CGA TAC    1200
165  Gln His Ile Glu Gln Leu Lys Gly Ser Ala Glu Gly Ser Ile Arg Tyr
166  385              390              395              400
167
168  CCC GCA TGG AAT GGG ATA ATA TCA CAG GAG GTA CTA GAC TAC CTT TCA    1248
169  Pro Ala Trp Asn Gly Ile Ile Ser Gln Glu Val Leu Asp Tyr Leu Ser
170              405              410              415
171
172  TCC TAC ATA AAT AGA CGC ATA TAAGTACGCA TTTAAGCATA AACACGCACT       1299
173  Ser Tyr Ile Asn Arg Arg Ile
174              420
175
176  ATGCCGTTCT TCTCATGTAT ATATATATAC AGGCAACACG CAGATATAGG TGCGACGTGA   1359
177
178  ACAGTGAGCT GTATGTGCGC A                                            1380
179
180
181  (2) INFORMATION FOR SEQ ID NO:2:
182
183          (i) SEQUENCE CHARACTERISTICS:
184                  (A) LENGTH: 423 amino acids
185                  (B) TYPE: amino acid
186                  (D) TOPOLOGY: linear
187
188          (ii) MOLECULE TYPE: protein
189
190          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
191
192  Met Pro Gln Phe Asp Ile Leu Cys Lys Thr Pro Pro Lys Val Leu Val
193   1               5               10              15
194
195  Arg Gln Phe Val Glu Arg Phe Glu Arg Pro Ser Gly Glu Lys Ile Ala
196          20              25              30
197
198  Leu Cys Ala Ala Glu Leu Thr Tyr Leu Cys Trp Met Ile Thr His Asn
199          35              40              45
200
201  Gly Thr Ala Ile Lys Arg Ala Thr Phe Met Ser Tyr Asn Thr Ile Ile
202      50              55              60
203
204  Ser Asn Ser Leu Ser Phe Asp Ile Val Asn Lys Ser Leu Gln Phe Lys
```

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/147,912*

INPUT SET: S811.raw

```
205    65                          70                          75                          80
206
207    Tyr Lys Thr Gln Lys Ala Thr Ile Leu Glu Ala Ser Leu Lys Lys Leu
208                85                          90                          95
209
210    Ile Pro Ala Trp Glu Phe Thr Ile Ile Pro Tyr Tyr Gly Gln Lys His
211                100                         105                         110
212
213    Gln Ser Asp Ile Thr Asp Ile Val Ser Ser Leu Gln Leu Gln Phe Glu
214                115                         120                         125
215
216    Ser Ser Glu Glu Ala Asp Lys Gly Asn Ser His Ser Lys Lys Met Leu
217        130                         135                         140
218
219    Lys Ala Leu Leu Ser Glu Gly Glu Ser Ile Trp Glu Ile Thr Glu Lys
220    145                         150                         155                         160
221
222    Ile Leu Asn Ser Phe Glu Tyr Thr Ser Arg Phe Thr Lys Thr Lys Thr
223                165                         170                         175
224
225    Leu Tyr Gln Phe Leu Phe Leu Ala Thr Phe Ile Asn Cys Gly Arg Phe
226                180                         185                         190
227
228    Ser Asp Ile Lys Asn Val Asp Pro Lys Ser Phe Lys Leu Val Gln Asn
229                195                         200                         205
230
231    Lys Tyr Leu Gly Val Ile Ile Gln Cys Leu Val Thr Glu Thr Lys Thr
232        210                         215                         220
233
234    Ser Val Ser Arg His Ile Tyr Phe Phe Ser Ala Arg Gly Arg Ile Asp
235    225                         230                         235                         240
236
237    Pro Leu Val Tyr Leu Asp Glu Phe Leu Arg Asn Ser Glu Pro Val Leu
238                245                         250                         255
239
240    Lys Arg Val Asn Arg Thr Gly Asn Ser Ser Ser Asn Lys Gln Glu Tyr
241                260                         265                         270
242
243    Gln Leu Leu Lys Asp Asn Leu Val Arg Ser Tyr Asn Lys Ala Leu Lys
244                275                         280                         285
245
246    Lys Asn Ala Pro Tyr Ser Ile Phe Ala Ile Lys Asn Gly Pro Lys Ser
247        290                         295                         300
248
249    His Ile Gly Arg His Leu Met Thr Ser Phe Leu Ser Met Lys Gly Leu
250    305                         310                         315                         320
251
252    Thr Glu Leu Thr Asn Val Val Gly Asn Trp Ser Asp Lys Arg Ala Ser
253                325                         330                         335
254
255    Ala Val Ala Arg Thr Thr Tyr Thr His Gln Ile Thr Ala Ile Pro Asp
```

```
256                 340              345              350
257
258    His Tyr Phe Ala Leu Val Ser Arg Tyr Tyr Ala Tyr Asp Pro Ile Ser
259            355              360              365
260
261    Lys Glu Met Ile Ala Leu Lys Asp Glu Thr Asn Pro Ile Glu Glu Trp
262            370              375              380
263
264    Gln His Ile Glu Gln Leu Lys Gly Ser Ala Glu Gly Ser Ile Arg Tyr
265    385              390              395              400
266
267    Pro Ala Trp Asn Gly Ile Ile Ser Gln Glu Val Leu Asp Tyr Leu Ser
268            405              410              415
269
270    Ser Tyr Ile Asn Arg Arg Ile
271            420
272
273    (2) INFORMATION FOR SEQ ID NO:3:
274
275        (i) SEQUENCE CHARACTERISTICS:
276            (A) LENGTH: 34 base pairs
277            (B) TYPE: nucleic acid
278            (C) STRANDEDNESS: unknown
279            (D) TOPOLOGY: unknown
280
281        (ii) MOLECULE TYPE: cDNA
282
283        (vi) ORIGINAL SOURCE:
284            (C) INDIVIDUAL ISOLATE: FLP recombination target site
285
286
287        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
288
289    GAAGTTCCTA TTCTCTAGAA AGTATAGGAA CTTC                              34
290
291    (2) INFORMATION FOR SEQ ID NO:4:
292
293        (i) SEQUENCE CHARACTERISTICS:
294            (A) LENGTH: 68 base pairs
295            (B) TYPE: nucleic acid
296            (C) STRANDEDNESS: unknown
297            (D) TOPOLOGY: unknown
298
299        (ii) MOLECULE TYPE: cDNA
300
301        (vi) ORIGINAL SOURCE:
302            (C) INDIVIDUAL ISOLATE: Synthetic oligonucleotide
303
304
305        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
306
```

```
307    GATCCCGGGC TACCATGGAG AAGTTCCTAT TCCGAAGTTC CTATTCTCTA GAAAGTATAG        60
308
309    GAACTTCA                                                                 68
310
```

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/147,912*

DATE: 01/07/94
TIME: 14:50:11

INPUT SET: S811.raw

Line   Error   Original Text

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION *US/08/147,912*

INPUT SET: S811.raw

< < THERE ARE NO ITEMS MISSING > >

PAGE: 1

SEQUENCE CORRECTION REPORT
PATENT APPLICATION *US/08/147,912*

DATE: 01/07/94
TIME: 14:50:12

INPUT SET: S811.raw

Line   Original Text   Corrected Text